

The modeling component of educational language semantics

Carreira Maria

Modern Language Association, USA

ABSTRACT

Article analyzes the existing approaches to the description of the component educational language semantics, proposed and justified the use of psycholinguistic databases, allowing applying statistical methods for analysis of semistructured semantic information. We consider the project of Tomsk State University for the development of such a framework: the instructions are presented samples of stimulus material, description of the data collection and analysis, preliminary findings. The possibilities of application of the developed database research practice. Domestic cognitive linguistics is widespread understanding of the concept as a complex mental inhomogeneous structure, which is often described by the field model (core, near and far periphery). Despite the fact that a set of specific elements of the structure and hierarchy can be defined differently, most researchers shared the view that it is konkretnoobraznye characteristics make up the core of the concept, or are among its key components: the figurative component image, concrete characteristics.

KEYWORDS: database, cognitive semantics, semantics of perceptive.

Perceptual information in the concept structure of connects directly shaped the concept of the component to perceptual sensations derived from the human body experience "more precise measurements of the most important concept - figurative, conceptual and value. Shaped side of the concept - it's visual, auditory, tactile, taste, smell perceived characteristics of objects, phenomena, events that appear in our minds, it is relevant features of practical knowledge"; on the same connection point Z.D. Popkin and I. Sternin, referring to the notion of a universal subject code (CPC) N.I. Zhink: "The presence in the concept shaped component is determined by the neuro-linguistic character of the universal subject code: sensual image encoding concept, forming one universal subject code". The importance of perceptual information con structure CEPT emphasizes A.A. Zalevskaya including for abstract con-

cepts: 'the meaning of each word as a unit idioleksikona, in principle, be reduced to a certain original sensuous image of the object (Visual, auditory, motor, etc., aktualizuemomu directly or through the mediation of verbal' transitions'), which should find expression in the detection of native language in the presence of the identifiable words (even with the most abstract value) of a certain degree of specificity and imagery".

Sometimes investigators go further and produce a structure within the shaped component. For example, Z.D. Popov and I. Sternin say perceptual image (combination of features formed during the direct perception of the external world fragment) and metaphorical or cognitive image (shaped sign formed metaphorical judgment result of an object or phenomenon). The basis of this classification is an indication of origin of the image (whether it is received through the senses or appeared as a result of formation of conceptual metaphor); In both cases, however, the image is made up of the same "mental material" of the same nature, so later in this paper the distinction between perceptual and figurative components is not done.

Existing methods for the study of perceptual component of language semantics.

Despite the fact that the importance of perceptual component in the structure of the concept is widely recognized, there is currently no existing method of permanently identifying perceptual semantics of language units.

Research informatsionnopolyatiynogo core simulation concept is carried out using conventional linguistic techniques (such as component analysis, context analysis dictionary definitions, work with phraseological pictures, etc.), and the result of this modeling is the set of semantic components forming the respective layer and concept expressed on the meta-language. These tools, however, are not as effective at modeling concept shaped component. The reason for this is seen in the fact that the very nature of perceptual information to others, this information is poorly amenable semiotization means of natural language, and because linguistic methods can be detected with difficulty. Although the language is closely included in the total cognitive system and can be viewed in two capacities - both are exposed to general psychophysiological laws, including perceptual as well as influencing the number of mental processes, including perception, - it is clear that not every kind of information contained in the human cognitive system, is equally easily accessible for analysis using proper linguistic methods, which makes researchers seek to find new methods of detection of perceptual information as an integral part of the concept. Although the language is closely included in the total cognitive system and can be viewed in two capacities - both are exposed to general psychophysiological laws, including perceptual as well as influencing the number of mental processes, including perception, - it is clear that not every kind of information contained in the human cognitive system, is equally easily accessible for analysis using proper linguistic methods, which makes researchers seek to find new methods of detection of perceptual information as an integral part of the concept. Although the language is closely included in the total cognitive system and can be viewed in two capacities - both are exposed to general psychophysiological laws, including perceptual as

well as influencing the number of mental processes, including perception [7.198-199], - it is clear that not every kind of information contained in the human cognitive system, is equally easily accessible for analysis using proper linguistic methods, which makes researchers seek to find new methods of detection of perceptual information as an integral part of the concept.

In the IA Sternin and M.J. Rosenfeld "Word and Image" offers two methods of figurative conceptual core component extraction: analysis of dictionary definitions and directed psycholinguistic experiment in which participants are invited to the following statement: "Describe everything you see, hear, feel, when the sound of each word of the experimental list". Below we briefly discussed the possibilities and limitations of each of the procedures.

When analyzing dictionary definitions authors show shaped component itself in interpreting the values in the examples accompanying the interpretation and reveal "gaps" in the interpretation that they believe to be filled descriptions sensory images (e.g., directional and indefinite pronouns - gdelibo, in chemlibo). The authors argue that the shaped component values expressed infinitives (eg, exit), which indicate the character, without naming him, as well as some abstract nouns semantics (mass, many, stream, string). In the case of specific concepts researcher often has to admit the definition of entirely "shaped". For example, leading to the word interpretation hand, the author's isolated shaped component 7 in the definition of the words 12 'each of the two upper to extremities of the human shoulder joint to the toes, as well as from the wrist to the fingertips'. Depending on the research position of a word or two, upper limb and the person also can easily be included in a "shaped" part of the interpretation. The authors point out that the definition may be found references to objects and phenomena, perceived by means of different senses (sight, hearing, taste, smell, and touch), emotionalnootsenochnye images and images of varying degrees of detail.

Directional associative experiment proposed by the authors allows to extract rich material for meaningful analysis; the task ("Describe everything you see, hear, feel, when the sound of each word of the experimental list") requires the respondents refer to their own perceptual experience related to the word, and does not impose any restrictions on the perceptual modality generated images (eg, visual, auditory modality, etc.).

Obviously, the method of analyzing dictionary definitions has several limitations: (1) it is based on a material which is not intended directly for the task (Semantics description pertseptivnoobraznogo component word); (2) it does not contain unequivocal allocation criteria perceptual component (such as its presence and boundaries in the text) and its classification, so that largely the problem is solved on the basis of research intuition; the researchers themselves note that "if ovneshlyaetsya sense-perception in a coherent text, or even in the same sentence, the description of the sensory image is open to all words of the sentence"; (3) because the vocabulary interpretation no intentional and systematic reflection on the perceptual information semantic component,

Proposed I.A. Sternin and M.J. Rosenfeld version of the association experiment provides the contrary, the material collected specifically for the task, and therefore allows the researcher to focus efforts on solving the main problem - the analysis of perceptual semantics component. Furthermore, by abstracting diskursivnoj specificity occurrence different perceptual components in the text, this method allows analysis of the relative importance of images of different modalities (visual, auditory, etc.) in the semantic structure of the word, provides material for perceptual content analysis of images, it answers the question " what kind of images emerge in the perception of the word? ".

Limitations of both techniques discussed above lies in the fact that these methods (1) lead to the inclusion in perceptual core each concept large number of weakly-ordered components of uncertain relationships between them, and (2) are time consuming, very difficult to collect manual method sufficiently representative sample in order to draw conclusions about the language and the cognitive system as a whole.

The following is a method of creating a psycholinguistic database, allowing to overcome these limitations by shifting perceptual semantics component in numerical form.

Using psycholinguistic databases in modern linguistics and cognitive science database technologies are actively used in modern linguistics

With the increasing amount of information needed to manage information flows effectively and quickly (to analyze structured data), the need to systematize and integrate data collected in different disciplines and using various methods in interdisciplinary research. Databases are used to working with the linguistic information of all kinds - phonological, lexical, morfemnoslovoobrazovatelnom information about the level of language, of the typological features of different languages. Particular important is the full-text databases, text body, methodological and practical importance which both linguistics in general and in particular cognitive linguistics currently intensively reflectors (see for example [11]). Active materials are also used parallel corpora,

An example of the database can serve as a frequency dictionary, such as "New frequency dictionary of Russian language" O.N. Lyashevsky and SA Sharov, established on the basis of the Russian National Corpus: each word is assigned a value (frequency), reflecting its position on some scale (from least to most frequency words frequency).

In the framework of psychology and psycholinguistics has become a tradition to use scales, designed for the quantitative measurement of semantic phenomena, ie Translation of non-numerical information in numerical - Ch semantic differential Osgood, a Likert scale, the essence of which consists in constructing semantic space (e.g., predetermined coordinates from 1 to 7) and placing the semantic object at a predetermined scale, i.e. at a certain point of the semantic space. For example, the respondent is asked to evaluate the degree of appearance of an attribute on a scale of 1 to 7, where 1 means very poor degree of display, and 7 - the greatest. Similarly, it can be converted into numerical form other types of semi-structured information. The framework of cognitive

linguistics, this method was supplemented by statistical analysis of a large sample of respondents of data in order to be able to judge not about the individual characteristics of perception and evaluation of semantic components and the general patterns that are valid for the average native speaker. Typically, in such a database is represented by two basic indicator for the set of values - average (mean), the average estimation reflecting the semantic object native speakers, and a standard deviation (standard deviation), which reflects the degree of scatter estimates. One of the first works to create such a psycholinguistic database was the work of A. Paivio, J. Yuille and S. Madigan «Concreteness, imagery, and meaningfulness values for 925 nouns», published in 1968, at which the average evaluation of collected metrics such as subjective concreteness / abstraction, imagery words (words ability to cause mental images of different modalities) subjective estimation ambiguity word. In the Keuleers and Balota «Megastudies, crowdsourcing, and large datasets in psycholinguistics: An overview of recent developments» provides an overview of Psycho and of linguistic databases that are used today in psycholinguistics, cognitive linguistics (including the corpus). The authors conclude that in the era of big data, these methods alter the presentation of the research process («This has opened new ways of doing psycholinguistics»).

Psycholinguistic scales that are used to create a database

In order to evaluate a variety of options, and is currently on the material in English and several other European languages, created a significant number of databases containing psycholinguistic assessment on various scales can be selected. Summarizing the information from different sources, is said to be the most important and actively used in psycholinguistic research practice types of scales.

I. A separate group of characteristics make up the parameters reflecting the psychophysiological and emotional human experience with these or other units of language:

- 1) imagery words (imageability);
- 2) subjective concreteness / abstraction;
- 3) physical interaction of the human body with the object (BOI, bodyobject interaction); the ability to manipulate objects using hands (Manipulability);

Feedback words modalities to individual perception (modality rating); assistant location in space on a scale; subjective assessment of speech (pleasantness); emotional words (emotionality).

In another group together characteristics that describe the operation of lexical units:

- 1) subjective understanding of age (subjective age of acquisition);
- 2) subjective frequency words (subjective frequency);
- 3) assessing whether word familiar (familiarity);
- 4) the availability of context1 (Context availability);
- 5) Assessment of the extent familiar with the idea expressed by a word (concept familiarity);
- 6) the number of associations of (number of associations);

7) the number of values ambiguity (number of meanings).

III. Many features are not fit to belong to a group, and there are situational in connection with the private research tasks. Here are some of them:

- 1) semantic transparency (semantic transparency);
- 2) ease of determination (ease of definition);
- 3) typicality (typicality, semantic typicality);
- 4) ease of word pronunciations (pronounceability).

IV. In addition to the subjective characteristics, researchers often include a database and objective, such as age understanding objective², The frequency of the objective (for example, can serve as already mentioned dictionary Lyashevskiy O.N. and Sharova S.A., based on the material of the Russian National Corpus). The objective, although obtained in include meaningfulness. During the experiment psycholinguistic characteristics, an indicator of the number of associations of data by study participants on the stimulus word per unit time (for example, 0.5 m or 1 m). This indicator reflects the ability of words easily and quickly evoke.

Often used in studies of a group of objective parameters that describe the form of the word (word surface form) - the number of letters, the number of sounds, the number of syllables, the stressed syllable number (for languages with a mobile accent), words with similar sounds, writing.

Translating semantic or other indicators in a numeric expression allows subsequent application of statistical methods to the received data - finding the arithmetic mean and standard deviation, analysis of the distribution count, the use of analysis of variance (ANOVA), using multiple scales can simultaneously apply correlation and regression analysis, factor analysis, cluster analysis. Statistical analysis of large data sets can shoot Research subjectivism and neutralize the unique characteristics of individual consciousness, the former is traditionally the subject of study psycholinguistics, finding, so the average values for the language units. It should be emphasized that this language elements and groups are units of such analysis is therefore

One of the first such works on Russian material is a dictionary "sense organs, emotions and adjectives of the Russian language", which contains information on the evaluation of communication Russian adjectives with different senses. Relevant, if somewhat different in the purposes of creation, is an electronic system "The noun and the object library of incentives and regulations for experimental studies" and "incentives Library: verbs and nouns" started to material research, presented in the [thirty].

Database design, modeling perceptual component of language semantics

The Tomsk State University, currently working on the creation a psycholinguistic database containing information on the relationship Russian nouns with the modalities of perception, ie, vision, hearing, taste, olfactory and tactile sensations, also additional information, such as for example the age of assimilation subjective word, subjective frequency of words, etc. At this point in the study involved more than 500 informants.

The selection of stimuli and evaluation parameters

To create the database were selected nouns of different semantic categories representing names: actions and processes (e.g., attack, jump, inhale, lifting, FAS The term); animals (e.g., Badger, bear, beetle, bird, cat, chameleon); body parts (e.g., crown, the ear, eyebrows, the face, foot, skull); buildings and parts thereof (e.g., a basement, a cellar, a mine, a canopy, a skyscraper); clothing and accessories (such as button, boots, earrings, headphones, toe ring); foods (e.g., mushroom, cranberries, apples, gooseberry, pumpkin, coconut); landscape elements, surface types (e.g., a mountain, a recess grass, pond, grass, asphalt); mental processes and objects (e.g., imagination, theory, plan, knowledge, dream, fantasy); feelings and emotions (e.g., resentment, love, jealousy, happiness, joy, confusion); physical sensations (e.g., chills, softness, pain, nausea, dryness); sounds (e.g., beep, jingle, squeak, rattle, cough, noise); of celestial objects and phenomena (e.g., star, moon, meteor, lightning, rainbow, dawn, cloud); tools and implements, actuated by means of fine motor skills (e.g., a key, a nail, a needle, a nut, clip, glass Sharpener); tools and implements driven by the motor via a large (e.g., a saw, a pan, a broom, a cup, a ball, a shovel, an ax); vehicles (e.g., train, car, coach, mo pad, wagon, aircraft, skate, tank); individual artifacts included in previous studies, and not related to category in this study (e.g., antenna curtains, trap, doormat, socket, flag); encoded in the database is given *GOVERNMENTAL* as a category of object.

Total nouns - 506. The database were both specific and abstract nouns that is not characteristic similar foreign databases, however this approach allows appears to compare concrete and abstract nouns together according to preset parameters. All words were assessed by the respondents on the following scale:

Imagery words (imageability) - the ease with which the word evokes images of the respondent objects, qualities, actions, etc. When evaluating the imagery of the words used by the following statement: The words vary in their ability to elicit mental images of things or phenomena. For example, when you hear or read the word "apple", it is likely that you can quickly and easily imagine an image of the subject. On the other hand, it is not so easy to imagine, for example, a "fact". Please vote the words in the list on a scale of 1 to 7, where 1 - the lowest score of imagery (the word with difficulty and slowly evokes images or cause at all), and 7 - the highest score of imagery (the word conjures up images quickly and easily). Keep in mind that this task can not be right or wrong answers, guided by their own feelings. Please do not break vote all the words for 1 session.

Communication with the individual words modalities of perception (modality rating) - sight, touch, hearing, taste, smell.

The next instruction is used in the evaluation of this parameter:

Different words may be related to different modalities of perception - sight, hearing, taste, smell and tactile sensations (sense of touch). Please vote the words in the list of their links with vision: 7 is the most powerful bond with sight words, 1 - the weakest link words with vision. Work as quickly as possible, without thinking long about every word. At the same time, try to be objective and focused. If you have the need arises, you can return to manual, and read it again, and then continues to evaluate nouns. Please do

not take breaks while working on this questionnaire, evaluate all the words for 1 session? For other modal most perception has been used a similar statement, in which the word view replaces the taste / smell / touch / hearing respectively.

After collecting primary information on the indicator words may be distributed over the modalities with which they are statistically significantly associated (i.e., values for the words to be transferred in a nominal scale) can be isolated modal (associated with only one modality) and polymodal (associated with different modalities) words.

The ability to manipulate the object by hand (manipulability). The next instruction is used in the evaluation of this parameter: To some items are performing, a person has to use his hands. For example, to cigarette fulfill its function («been smoked»), the use of hands necessary. On the other hand, a person will never interact with their hands from the volcano to the function performed («erupted»). Please vote the words in the list, depending on how you need to use human hands items to fulfill their typical functions: 7 - the use of arms must always be 1 - hands are never used to interact with the subject. Some objects can act in different capacities: as an oyster can be an ingredient dishes (and then people use their hands, to Oyster "was prepared"); On the other hand, it is a living being, and to oyster "lived" the people do not need to interact with it by hand. In such cases, guided by the option that first comes to mind.

Location of the referent in the space on the scale top-down - assessment of how high or low relative to the observer is the referent word. This feature was used in this project, not only for the evaluation of specific nouns, which assistants can be perceived through the senses, but also for the abstract vocabulary. Some of the objects and phenomena of reality are usually located above or below in the space: The following statement was used. Please rate how high or low is the subject, using a scale from 1 to 7, where 1 - "very low", and 7 - "very high". Work as quickly as possible, without thinking long about every word. At the same time, try to be objective and focused. If you have the need arises, you can go back to the manual and read it again, and then continue to evaluate nouns. Please do not take breaks while working on this questionnaire, evaluate all the words for session?

Subjective age of understanding (subjective age of acquisition) - an indicator reflecting the age at which the respondents, in their opinion, have learned collateral to the word. Different words, we learn at different ages: the following statement is used in the evaluation of this parameter. Please put next to each word in the list of the number that corresponds to the age (in years), when you have learned the word. Please use the numbers from 0 to 15. Put a dash if the word is unfamiliar to you. Work as quickly as possible, without thinking long about every word. At the same time, try to be objective and focused. If you have the need arises, you can return to manual, and read it again, and then continues to evaluate nouns. Please do not take breaks while working on this questionnaire.

Subjective frequency words (subjective frequency) - the extent to which part of the respondents, in their opinion, meet with a particular word or use it. This parameter was used follows

Blowing instruction: Different words are used in a speech at different rates. Please rate how often you meet with every word, using a scale from 1 to 7, where 1 - "very rare", and 7 - "very often". Work as quickly as possible, without thinking long about every word. At the same time, try to be objective and focused. If you have the need arises, you can return to manual, and read it again, and then continues to evaluate nouns. Please do not take breaks while working on this questionnaire, evaluate all the words for 1 session?

The selection of data and evaluation parameters semantic groups conducted in accordance with the latest neurocognitive studies. In particular, it was shown that treatment of words that refer to different perceptual modalities, leads to activation of various neural patterns in the brain (see. Overview in), in the same way varies and word processing belonging to a different semantic groups, up to subtle differences names of actions performed by means of an arm or leg; Patients with neurophysiologic disorders (such as Alzheimer's disease) sometimes exhibit loss of knowledge of some categories (plants, animals), while maintaining the other (tools, furniture), which is likely to be associated with various forms of human interaction with the representatives of the various categories and , Consequently, with different perceptual experiences stored in consciousness. The importance of perceptual modalities like lo component mantic structure, its psychological reality demonstrated in behavioral experiments measuring the response time, and even idiomatic expressions where components are used in a metaphorical sense, such as in terms helping hand. In general, these studies are consistent with the fundamental position of anthropocentrism language at all levels, which is taken in the national cognitive linguistics (e.g.), developing the idea in line with the theory of embodied cognition (embodied cognition, see. for example).

Data collection

Respondents filled out questionnaires in paper form, then converted into an electronic form (use the Microsoft Office Excel). Each questionnaire was printed on an A4 sheet and contains instructions and 60 randomly selected words from the general list. Each experiment participant received at the hands of five such profiles, all profiles - with different parameters of evaluation; not one word was repeated twice for one participant. On the experiment participants were going to the following demographic information: gender, age, level of education (student vs. student vs. higher education), education direction (humanitarian vs. vs. natural-technical). Also gathered information on the respondent's own language; of those respondents who reported not as a native Russian language, they were excluded from further analysis. If desired, participants could pass the experiment again, with new words and / or a scale for the assessment; in which case they were recorded as new members. If possible, a uniform distribution profiles observed among the respondents of both sexes.

In order to achieve the statistical reliability of the results for each word for each scale has been collected about 30 assessments.

Processing the data collected, the possibility of using the results

Collected during the experiment psycholinguistic assessment of the study participants were averaged, we found the standard deviation (use the statistical).

The minimum and maximum values - the largest and smallest values in the sample obtained after an average of 30 estimates, data on the members of each word. For example, in power communication with tactile sensations maximum value (5.93 of 7) received word cat, i.e. in this sample is the word most strongly linked to the respondents with tactile sensations. By the same parameter minimum value (1.14 of 7) received word manner, i.e. it is less likely due to the tactile sensations. Another example is the word raspberries and carrot receiving the maximum values in the degree of imagery (7 of 7), i.e. those words faster and easier to produce images in the mind. The minimum value (1.93 of 7) in the degree of imagery received word assumption, i.e. Respondents hardly form an image in the perception of the semantics of the word.

Descriptive statistics: Vis - the binding force of speech with the visual modality (1 to 7); Audi - the power of words with auditory communication modality (1 to 7); Of - force communication modality words with olfactory (smell) (1 to 7); Gus - the power of speech communication with a modality of taste (1 to 7); Hap - force due to the tactile modality of words (1 to 7); Image - degree imagery words (1 to 7); Man - able to interact with referent by means of hands (1 to 7); Space - localization referent in the space (1 - bottom 7 - top); a subjective age assimilation of the word (in years); Mean - mean value, SD - standard deviation, Min - minimum value, Max - maximum value (statistic found on material 506 values).

Standard deviation (standard deviation, SD) indicates how much the values in the sample are scattered around the mean. For example, it can be seen that the standard deviation for connection with the visual modality words less than with flavoring (1.03 and 1.39 of 7 respectively). This means that the values on the scale due to the visual modality more "tightly" centered around the mean value (3.98 in this case), i.e. most of the sample cells is similar for this parameter, while the values of strength due to the taste modality is more strongly scattered around the mean (2.17) and, as can be seen in Fig. 3, even here are the two clearly distinguishable groups of words - one (small in number) with a strong bond with the sense of taste (in the upper part of the chart) and the second,

The average value is obtained by a simple arithmetic averaging of the sampling values (in these case 506 values) and reflects the overall "center of gravity" of the sample for a given parameter. As can be seen from those shown in Fig. 1 data, the mean value of the sample on a scale communication words vision is the highest among the five modalities (3.98), followed by the sense of touch is (3.39), then - the ear (2.32); the weakest bonding force sample words with taste (2.71) and smell (2.32). These patterns correspond to previously detected within perceptual psychology man receives the greatest amount of information through sight, followed by hearing and sense of touch, the smallest role olfactory and taste sensations.

The analysis shows that the semantics of natural language is isomorphic to the structure of human perception. It is also found that the average value of the scale imagery (5,6) is substantially above the middle of the scale (3.5) and approaches the upper

border (7), which correlated with the results of research on foreign material other languages and confirms that understanding natural language semantics It relies heavily on the images of different nature, component.

Similarly it can be calculated descriptive statistics for individual groups of words selected for various reasons (formal semantic). For example, in Fig. 4 presents descriptive statistics on the scale of imagery for each of the semantic categories of words in the database. Analysis shows that the different semantic categories of words differ significantly from the average value of imagery, more clearly shown.

Group	N	Img. Mean	Img_SD
Action	thirty	4.03	1.05
Animal	76	6.58	0.27
Body_Part	14	6.29	0.44
Building	8	6.1	0.51
Clothes	28	6.43	0.47
Food	51	6.59	0.45
Ground	25	5.84	0.9
Intelligence	29	3.07	0.67
Object	15	6.23	0.42
Sense_Emotion	45	3.6	0.59
Sense_Phys	21	3.74	0.7
Sound	17	3.74	0.67
Space	18	6.09	0.98
Tool_power_grip	34	6.29	0.8
Tool_precise_grip	48	6.38	0.49
Transport	47	6.3	0.62
All_groups	506	5.6	1.4

Descriptive statistics for imagery (image ability) by semantic categories:

Group - semantic category; N - the number of units in the category; Img. Mean - the average value on the scale of the imagery; Image - standard deviation on the scale of the imagery; Action - names of actions and processes; Animal - Animal names, Body Part - names of body parts; Building - names of buildings and parts thereof; Clothes - names of items of clothing and accessories; Food - food items; Ground - here landscape elements, surface types; Intelligence - the name of the mental processes and facilities; Object - the name of the individual artifacts (not included in other categories); Sense Emotion - names of feelings and emotions; Sense_Phys - the name of the physical sensations; Sound - the name sounds; Space - the name of celestial objects and phenomena; Tool_power_grip - name tools guns actuated via gross motor; Tool_precise_grip -

names of tools and implements driven by using fine motor skills; Transport - the name of the vehicles; All_Groups - totals for the entire sample

A graphical representation of the descriptive statistics for the imagery (image ability) by semantic categories. Average values for each category are shown as squares; vertical lines represent the standard deviation in the deferred more and less side. Designate categories cm. In the explanation of the example, it can be concluded that the names of actions, mental states and objects, physical sensations and sounds, on average are significantly less imagery than nouns other semantic categories (it should, however, be noted that this preliminary conclusion reached on the basis of subjective evaluation graphics; to assert the existence of significant differences between categories on the basis implying for any use appropriate statistical tests, such as Kruskal - Wallis).

Fig. 6. Fragment database: WORD - word; Category - semantic group (category designations refer to the explanation for); Vis - average value of the coupling strength with the visual modality words (1 to 7); Audi - mean value of the force due to the auditory modality words (1 to 7); Of - average force communication modality words with olfactory (smell) (1 to 7); Gus - average value of the coupling strength with the word taste modality (1 to 7); Hap - the average force of speech communication with modality (1 to 7); Image - average value of the imagery words (1 to 7); Man - average value on a scale manipulability (to react with referent by hand) (1 to 7); Space - localization referent in the space (1 - bottom 7 - top); Ana - a subjective age assimilation of the word (in years the presented data base allows to apply to the variable semantic statistical methods, both directly comparing the values of groups of words (allocated on a particular base, such semantic, grammar and so on), using ANOVA analysis of variance, and comparing the values at different scales between a (i.e., to find the correlation between the scales). Here are a number of further installed on the collected database correlations psycholinguistic features of words.

Thus, for example, carried out correlation analysis of data collected in the database it possible to establish a strong correlation between such parameters as the relationship keywords and vision imagery words ($r = 0,73$), which can be interpreted as evidence that it is visual sensations contribute most to the imagery of the words; For example, the word fireworks, the tiger, apple, get a high score on a scale imagery, also received high marks in strength due to the vision, while the word assumption, conjecture, meaning received low scores on both scales. On the other hand, although these parameters are related to each other, they are not identical: it is possible to find words with a high degree of imagery, but a weak force due to visual sensations (breath rattling, horn).

It is also highly correlated with each other taste and olfactory sensations reflected in the semantic structure of the word ($r = 0,74$), which corresponds to the features of human solid interaction with the outside world as a rule, a person knows the smell of the objects that is tasting the (food) therefore, in his mind, olfactory and gustatory information is often related to each other. For example, set a high degree of communication simultaneously with the olfactory and taste sensations have words such as raspberry, orange, cherry; at the same time there are words with considerable force due to the ol-

factory sensations that are not related with the flavored - damp, dog, lawn. The inverse ratio (nouns related to taste and non-olfactory sensations) observed in fewer examples: words,

Furthermore, it found many other significant but less strong correlation between the individual parameters. Thus, for example, revealed a negative correlation between age and assimilation imagery words (-0.56), i.e. the earlier a person learns the word, the more imagery it has, which can be explained as a proper cognitive mechanisms (sequential development of thinking from to abstract) and discursive environments, which is a person in the course of their learning and maturation (development terminology, abstract vocabulary in the process of study); both aspects, of course, closely linked.

Based on a database also possible to study the relationship of imagery with different modalities of perception: if between the imagery and communication with visual modality there is a strong correlation, as mentioned above, between imagery and auditory modality significant correlations were found. It can be concluded that the auditory sensations "visualize" more complicated than with feelings of other modalities; moderately correlated with the value of imagery on a scale of tactile sensations, weak correlation observed between the imagery and the taste and olfactory sensations.

Databases of this type are important for modern domestic psychosocial and cognitive linguistics in the context of the treatment of the latter to hardware experimental studies - measurement of the reaction time (. See, e.g.), research using technology (see, eg.). For this kind of experimental studies are usually required stimulus material, are controlled by various psycholinguistic parameters, and the first stage of the study is the selection of the material, which greatly facilitated by the availability of ready databases, which contain information about the most common and popular psycholinguistic characteristics of different incentives kind (lexical, phrase logical units, metaphors, etc.).

Discussion of the possible methods of restrictions.

The most obvious limitation in terms of the usual linguistic description is the numerical format of the semantics, which does not allow talking about substantive differences between points on a single scale. This limitation is a natural continuation of the advantages of the method: removing the semantic heterogeneity in mathematical modeling, it enables the mass compared with each other linguistic units by selected criteria, applying statistical procedures to semantic variables. This method can be considered as an additional and relative to the direction associative experiments reported in I.A. Stern in and M.J. Rosenfeld "Word and Image" (see. Above), where the association experiment reveals the substantive content of the perceptual component,

Another possible observation is that in the central unit of semantic studies should be lexical semantic variant, a separate meaning of the word, while in the questionnaire survey participants work with the word. This limitation is common to all psycholinguistic techniques, in which the words are presented out of context (including the association experiment, discussed above): a researcher clearly defines the word with the help of context and then cannot eliminate the influence of the context or the word is presented in isolation, and then you cannot say with full confidence, with which it works LSV

study participants. This limitation is partly overcome by the fact that different LSV have different frequencies, the closest to the actual layer of consciousness. In addition, the inclusion of such an index database as standard deviation allows to identify the degree of difference between the estimates and identify individual members, so what words have caused the greatest differences in the interpretation of (possibly due to different actualization LSV).

Conclusion

The problem of describing the semantics perceptual component is part of the fundamental problems of establishing relations language and perception, traditional linguistic (reviewed in, And therefore unlikely to be the method can be solved; productive to use different methods to achieve a private research purposes. This paper presents a method of creating a psycholinguistic databases, allowing translate semi structured and / or hard to describe the information on the semantics of the perceptual component in numerical form, which makes it possible to use statistical procedures when dealing with semantics. Perceptual information word in the semantic structure is presented as mean values on a scale with five connection words perception modalities (vision, hearing, taste, smell, touch); moreover, found the average values of indicators such as the imagery of words (the ability to quickly and easily evokes images in the mind), the ability to manipulate by hand, the localization of the referent in the space on the vertical axis, the subjective age of mastering words. Collected data base can be used both as an independent semantics analysis tool and as a source of stimuli for psycholinguistic cognitive and experimental studies.

REFERENCES

- Vorkachev, S.G. (2004) Happiness as a linguocultural concept. Moscow: Gnosis.
- Inciting, MV (2004) Expanded theses on concepts. 1. pp. 53-64.
- Popover, ZD & Stern in, IA (2007) Cognitive linguistics. Moscow: AST: Vostok Zapad.
- Boldyrev, NN (2001) Cognitive semantics: A course of lectures on English philology. 2nd ed. Tambov: Tambov State University.
- Karasik, V. (2002) Language circle: Person, concepts, discourse. Volgograd: Permian.
- Zalevskaya, A.A. (1990) The word in the lexicon of the person: A psycholinguistic research. Voronezh: Voronezh State University.
- Rezanova, Z.I. (2010) Cognitive linguistics in the paradigms of linguistic functionalism and integral concepts of cognition. Tomsk State University Journal. 334. pp. 195-199. (In Russian).
- Sternin, I.A. & Rozenfel'd, M.Ya. (2008) Word and image. Voronezh: Istoki.
- Kuznetsov, S.A. (ed.) (1998) Great Dictionary of the Russian Language. St. Petersburg: Norint.
- Mishankina, N.A. (2013) Databases in linguistic research. Voprosy leksikografii - Russian Journal of Lexicology. 13). pp. 25-34. (In Russian).
- Zabotkina, V.I. (Ed) (2015) The methods of cognitive analysis of the semantics of the words: computercorpus approach. Moscow: Yazyki slavyanskoy kul'tury. [Online] Available from: http://www.rfbr.ru/rffi/ru/books/o_1924029#1. (Accessed: 16th July 2016).

- Nagel, O. (2014) Investigating Russian Derivational Suffix - yaka: Russian Parallel Corpus Study. *Procedia Social and Behavioral Sciences*, 154. pp. 122-129. DOI: 10.1016 / j.sbspro.2014.10.123
- Lyashevskaya, ON & Sharov, SA (2009) Frequency Dictionary of Modern Russian language (on materials of the Russian National Corpus. Moscow: Azbukovnik. [Online] Available from: <http://dict.ruslang.ru/freq.php>. (Accessed: 16th July 2016).
- Paivio, A., Yuille, JC & Madigan, SA (1968) Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology. Monograph Supplement*. 76 (1: 2). pp. 1- 25.
- Keuleers, E. & Balota, DA (2015) Megastudies, crowdsourcing, and large datasets in psycholinguistics: An overview of recent developments. *The Quarterly Journal of Experimental Psychology*. 68: 8. pp. 1457-1468.
- Juhasz, BJ, Lai, YH & Woodcock, ML (2015) A database of English compound 629 words: ratings of familiarity, lexeme meaning dominance, semantic transparency, age of acquisition, imageability, and sensory experience. *Behavior Research Methods*. 47 (4). pp. 1004-1019.
- Altarriba, J., Bauer, LM & Benvenuto, C. (1999) Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, and Computers*. 31 (4). pp. 578-602.
- Bennett, SDR et al. (2011) Imageability and bodyobject interaction ratings for 599 multisyllabic nouns. *Behavior Research Methods*. 43. pp. 1100-1109.
- Brysbaert, M., Warriner, AB & Kuperman, V. (2014) Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*. 46. pp. 904-911.
- Tillotson, SM, Siakaluk, PD & Pexman, PM (2008) Body-object interaction ratings for 1618 monosyllabic nouns. *Behavior Research Methods*. 40 (4). pp. 1075-1078.
- MorenoMartínez, FJ, Montoro, PR & RodríguezRojo, IC (2014) Spanish norms for age of acquisition, concept familiarity, lexical frequency, manipulability, typicality, and other variables for 820 words from 14 living / nonliving concepts. *Behavior Research Methods*. 46 (4). pp. 1088-1097.
- Lynott, D. & Connell, L. (2009) Modality exclusivity norms for 423 object properties. *Behavior Research Methods*. 41. pp. 558-564.
- Lynott, D. & Connell, L. (2013) Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods*. 45. pp. 516-526. DOI: 10.3758 / s1342801202670
- Clark, JM & Paivio, A. (2004) Extensions of the Paivio, Yuille, and Madigan (1968) norms. *Behavior Research Methods, Instruments, & Computers*. 36 (3). pp. 371-383. DOI: 10.3758 / BF03195584
- Schröder, A. Et al. (2012) German norms for semantic typicality, age of acquisition, and concept familiarity. *Behavior Research Methods*. 44. pp. 380-394. DOI: 10.3758 / s134280110164y
- Russian National Corpus. Available from: <http://www.ruscorpora.ru>. (Accessed: 16th July 2016). (In Russian).
- Kolbeneva, MG & Aleksandrov, Yu.I. (2010) Senses, emotions and adjectives of the Russian language: Linguopsychological dictionary. Moscow: Yazyki slavyanskikh kul'tur.

- The Noun and the Object: Library of Stimuli and Regulations for Experimental Research. [Online] Available from: <http://nounobject.ru>. (Accessed: 16th July 2016). (In Russian).
- The Library of Stimuli. Available from: <http://stimdb.ru>. (Accessed: 16th July 2016). (In Russian).
- Akinina, Y. et al. (2015) Russian normative data for 375 action pictures and verbs. *Behavior Research Methods*. 47 (3). pp. 691-707.
- Pulvermüller, F. (2012) Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge. *Journal of Neurolinguistics*. 25: 5. pp. 423-459. DOI: 10.1016 / j.jneuroling.2011.03.004
- Kiefer, M. & Pulvermüller, F. (2012) Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*. 48 (7). pp. 805-825. DOI: 10.1016 / j.cortex.2011.04.006
- Connell, L., Lynott, D. & Dreyer, F. (2012) A Functional Role for Modality Specific Perceptual Systems in Conceptual Representations. *PLoS ONE*. 7 (3): e33321. DOI: 10.1371 / journal.pone.0033321
- Pecher, D., Zeelenberg, R. & Barsalou, L. (2003) Verifying DifferentModality Properties for Concepts Produces Switching Costs. *Psychological Science*. 14: 2. pp. 119-124. DOI: 10.1111 / 14679280.t01101429
- Gradinarova, G., Kanchev, P. & Janyan, A. (2014) Give Me Time to Picture That: Effects of Time and Imageability on EffectorSpecific Motor Activation in Idiom Processing. *The Russian Journal of Cognitive Science*. 1 (1-2). pp. 22-30.
- Rakhilina, E.V. (2008) Kognitivnyy analiz predmetnykh imen: semantika i sochetaemost' [Cognitive analysis of subject names: semantics and compatibility]. Moscow: Russkie slovari.
- Barsalou, L.W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences*. 22 (4). pp. 577-609; discussion: pp. 610-660. DOI: 10.1017 / S0140525X99002149
- Velichkovskiy, B.M., Zinchenko, V.P. & Luriya, A.R. (1973) *Psikhologiya vospriyatiya* [The psychology of perception]. Moscow: Moscow State University. Available from: <http://psychlib.ru/mgppu/vpv/VPV001.HTM>
- Rezanova, Z., Nekrasova, E. & Shilyaev, K. (2014) Gendermarked metaphors: influence of grammatical gender and animateness on referential choice of metaphorical name of the person in the Russian language. *Procedia - Social and Behavioral Sciences*. 215. pp. 273-278.
- Novikova, EG, Janyan, A. & Tsaregorodtseva, OV (2015) Metaphorical Salience in Artistic Text Processing: Evidence From Eye Movement. *Perception*. 44 (89). pp. 1098-1102. DOI: 10.1177 / 0301006615596692
- Kubryakova, ES (2004) Language and knowledge: On the way of learning the language: Parts of speech with a cognitive point of view. The role of language in the knowledge of the world. Moscow: